Dual concentration-dependent activity of thyroglobulin type-1 domain of testican: specific inhibitor and substrate of cathepsin L

Meh, Primož ; Pavšič, Miha ; Turk, Vito ; Baici, Antonio ; Lenarčič, Brigita

Abstract: The thyroglobulin type-1 (Tg-1) domain is a protein module that occurs in a variety of secreted and membrane proteins and is recognised as a potent inhibitor of cysteine peptidases. We present here some properties of the Tg-1 domain of human testican, a modularly organised proteoglycan secreted mainly by brain cells, the exact in vivo function of which is not yet clear. The domain was prepared as a recombinant protein in a Pichia pastoris expression system and its activity was demonstrated by specific and selective inhibition of cathepsin L (K i=0.14nM). Interaction at high enzyme and inhibitor concentrations resulted in degradation of the domain by cathepsin L, which was not observed under conditions used for the determination of kinetic parameters. No inhibitory activity could be detected for cathepsin K, but it exhibited a very similar degradation pattern. Homology modelling provided a good explanation for the different behaviour observed with the two enzymes. Firstly, the steric fit between the interfaces of testican domain and cathepsin L is stabilised by numerous favourable forces, while no such interactions are evident in the complex with cathepsin K, and repulsive interactions even prevent access of the domain to the active site of papain. Secondly, the prolonged first loop of the domain occupies a position near the catalytic cysteine residue in a more substrate-like manner, enabling cleavage of the Gly22-Ala23 bond

DOI: https://doi.org/10.1515/bc.2005.010

Biol. Chem., Vol. 386, pp. 75–83, January 2005 • Copyright © by Walter de Gruyter • Berlin • New York. DOI 10.1515/BC.2005.010

Dual concentration-dependent activity of thyroglobulin type-1 domain of testican: specific inhibitor and substrate of cathepsin L

Primož Meh[1], Miha Pavšič[1], Vito Turk[1], Antonio Baici[2] and Brigita Lenarčič[1,3,*]

[1] Department of Biochemistry and Molecular Biology, Jozef Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia

[2] Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

[3] Department of Chemistry and Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Jamova 39, SI-1000 Ljubljana, Slovenia

* Corresponding author
e-mail: brigita.lenarcic@ijs.si

Abstract

The thyroglobulin type-1 (Tg-1) domain is a protein module that occurs in a variety of secreted and membrane proteins and is recognised as a potent inhibitor of cysteine peptidases. We present here some properties of the Tg-1 domain of human testican, a modularly organised proteoglycan secreted mainly by brain cells, the exact *in vivo* function of which is not yet clear. The domain was prepared as a recombinant protein in a *Pichia pastoris* expression system and its activity was demonstrated by specific and selective inhibition of cathepsin L (K_i=0.14 nM). Interaction at high enzyme and inhibitor concentrations resulted in degradation of the domain by cathepsin L, which was not observed under conditions used for the determination of kinetic parameters. No inhibitory activity could be detected for cathepsin K, but it exhibited a very similar degradation pattern. Homology modelling provided a good explanation for the different behaviour observed with the two enzymes. Firstly, the steric fit between the interfaces of testican domain and cathepsin L is stabilised by numerous favourable forces, while no such interactions are evident in the complex with cathepsin K, and repulsive interactions even prevent access of the domain to the active site of papain. Secondly, the prolonged first loop of the domain occupies a position near the catalytic cysteine residue in a more substrate-like manner, enabling cleavage of the Gly[22]-Ala[23] bond.

Keywords: cysteine peptidases; extracellular proteolysis; inhibition kinetics; proteoglycans; testican; thyropin.

Introduction

Thyroglobulin type-1 (Tg-1) domains are recognised as unique sequence motifs of 65–80 residues. They occur in a variety of modularly organised proteins that are either secreted or located in membranes (Molina et al., 1996; Lenarčič and Bevec, 1998). A few years ago, it was established that some proteins containing a Tg-1 module function as inhibitors of the papain-related cysteine peptidases. So far, five proteins have been characterised as true inhibitors, the p41 invariant chain fragment (Bevec et al., 1996), chum salmon egg cysteine peptidase inhibitor (Yamashita and Konagaya, 1996), equistatin (Lenarčič et al., 1997), saxiphilin (Lenarčič et al., 2000) and testican (Bocock et al., 2003). All these inhibitors differ significantly in their selectivity for target enzymes. For example, the p41 fragment can discriminate between two closely related enzymes, cathepsin L and cathepsin S, inhibiting only the former (Bevec et al., 1996). Only one of the two Tg-1 domains present in saxiphilin inhibits cathepsins L or B (Lenarčič et al., 2000) and, in the case of the three equistatin domains, the first one exhibits broad specificity against papain-like cysteine peptidases, while the subsequent two are non-inhibitory (Lenarčič and Turk, 1999; Galeša et al., 2003). Furthermore, the second Tg-1 domain of equistatin is unique in inhibiting the aspartic peptidase cathepsin D (Galeša et al., 2003). The general inhibitory mechanism of a Tg-1 domain was revealed by the crystal structure of the p41 fragment in complex with cathepsin L (Gunčar et al., 1999). The wedge shape and the three-loop arrangement of the p41 fragment interacts with the active site cleft in a non-specific manner, whereas the additional contacts between regions of the inhibitor with the right domain of the enzyme are supposed to be crucial for the specificity. The structure of the p41 fragment is stabilised by three disulfide bridges and is also well defined in the absence of enzyme (Chiva et al., 2003).

Testican is a proteoglycan expressed predominantly in brain (Bonnet et al., 1996). At the genetic level, two additional forms of testican have been identified, testican-2 (Nagase et al., 1996; Vannahme et al., 1999) and testican-3 (Nakada et al., 2001). The characteristic feature of all variants is their modular organisation. The signal sequence on the N-terminal end is followed by a short part unique to testican, the follistatin domain, which also includes a Kazal domain, the calcium-binding domain, by the thyroglobulin type-1 domain (Tg-1), and finally by a short C-terminal sequence where two glycosaminoglycan attachment sites are located. The latter modification is suggestive of its involvement in cell adhesion processes and extracellular matrix organisation (Hartmann and Maurer, 2001). Lysosomal peptidases are known to be secreted into the peri/extracellular space, where they take part in remodelling or pathological degradation of extracellular matrix proteins in processes such as osteoarthritis, angiogenesis and inflammatory response (Baici

et al., 1995; Chapman et al., 1997; McGrath, 1999). Recently, a model of the involvement of the p41 invariant chain Tg-1 domain in the regulation of extracellular cathepsin L activity was presented (Fiebiger et al., 2002). It was shown that a stable complex of cathepsin L/p41 fragment is formed within antigen presenting cells and is afterwards secreted into the extracellular milieu, where the Tg-1 domain of p41 most probably acts to protect the enzyme and is displaced by a physiological substrate when necessary.

In the present study we characterised the Tg-1 domain, a 7-kDa fragment of the 49-kDa full-length human testican, which shows significant sequence similarity to the Tg-1 domains of the p41 invariant chain and equistatin, for which the specificities against different peptidases have already been determined in detail (Bevec et al., 1996; Lenarčič and Turk, 1999; Lenarčič et al., 2000). Using the expression system of *Pichia pastoris*, we produced the Tg-1 domain in an apparently native form. Molecular modelling and functional analyses revealed new features about the interaction of testican Tg-1 domain with its target papain-like enzymes.

Results and discussion

Preparation of recombinant Tg-1 domain of testican

P. pastoris proved to be a highly effective heterologous system for generating functional Tg-1 domains of equistatin in high yield (Galeša et al., 2003). In the present study, we utilised the same system to produce the Tg-1 domain of human testican. The cDNA construct for Tg-1 domain coding for 70 amino acids (Gly^{310}–Glu^{379}, testican unprocessed precursor counting) was introduced into the *P. pastoris* GS115 strain and the level of the expressed protein in media was followed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (data not shown). The domain was purified by three chromatographic steps and at the end a single band on SDS-PAGE corresponding to a molecular mass of approximately 8 kDa was obtained (data not shown). In the final purity assessment by N-terminal sequencing, we detected three slightly different populations of domains. In approximately 60% of the molecules two Glu-Ala repeats (out of three) were missing at the N-terminus, in 25% of the molecules one Glu-Ala repeat had been removed, and 15% of the molecules had the sequence of the intact Tg-1 domain. Processing of α-MF propeptide from the recombinant protein by the Kex2 endopeptidase appeared to be efficient, while the efficiency of Ste13 in removing the Glu-Ala repeats was not absolute. Similarly, it was observed that the two Glu-Ala repeats begin the sequences of previously reported Tg-1 domains of equistatin expressed in *P. pastoris* (Galeša et al., 2003), confirming the fact that the precise action of Ste13 depends on the surrounding amino acid sequence and on the tertiary structure of the protein of interest (Brake et al., 1984; Cereghino and Cregg, 2000). A yield of 10 mg of pure, active testican Tg-1 domain was typically obtained from 1 l of culture medium.

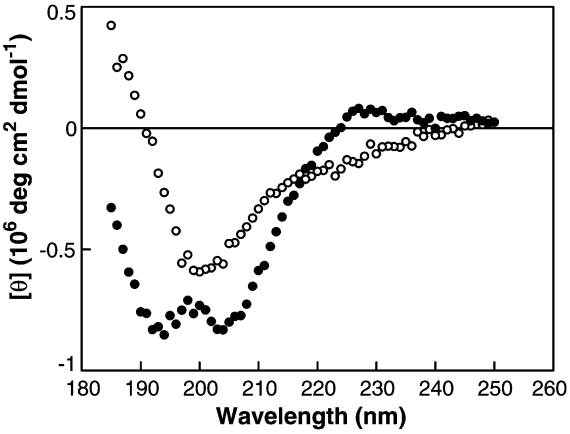


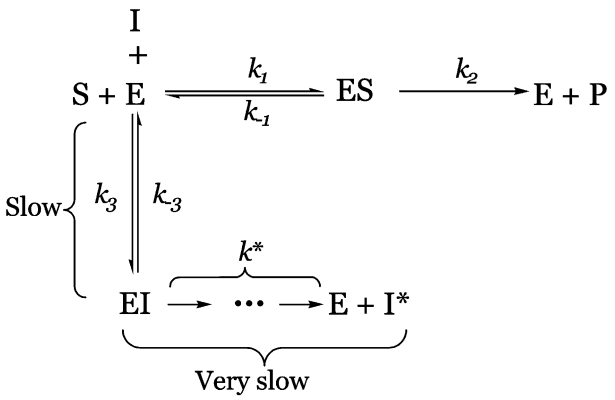
Figure 1 CD spectra of the testican Tg-1 domain under normal and denaturing conditions.
CD spectra were recorded using a protein concentration of 0.1 mg/ml in 10 mM sodium phosphate buffer, pH 5.60. (●) Testican Tg-1 domain measured at 4°C; and (○) testican Tg-1 domain in the presence of 12.5 μM DTT, measured at 88°C.

CD spectrum of testican Tg-1 domain

The CD spectrum of the Tg-1 domain of testican is characterised by two peaks, a positive one at 230 nm and an intense double trough at 204 nm (Figure 1). The spectrum differs slightly from the spectra of the individual Tg-1 domains of equistatin (Galeša et al., 2003) in two respects, by having a pronounced double trough and by the fact that both peaks are shifted to slightly higher wavelengths. As a comparison, the first domain of equistatin has negative peaks at 200, 194 and 198 nm, while the positive peak, which reflects the aromatic contribution (Woody, 1994) is present only in the second domain at 220 nm. A similar strongly positive band at 226 nm has also been observed with synthetic p41icf Tg-1 domain (Chiva et al., 2003). The small differences observed in the CD spectra are probably a consequence of slightly different backbone conformations of individual domains, which might in turn correspond to different specificities. On the other hand, it appears that the general features of all known Tg-1 domain spectra are characteristic of the thyropin fold. It very much resembles spectra of a group of β-proteins in which the α-helix content is low (Koepf et al., 1999; Kidrič et al., 2002). When the testican domain was thermally unfolded, the spectrum exhibited the characteristics of a random coil (Figure 1).

Inhibition of cathepsin L by the Tg-1 domain of testican

An important characteristic of thyroglobulin type-1 domains is their selectivity in inhibiting different papain-like enzymes (Bevec et al., 1996; Lenarčič et al., 1997, 2000). To investigate the inhibitory capacity of the testican Tg-1 domain we considered the human lysosomal cysteine peptidases cathepsins B, K, L, S, and X, and the plant cysteine peptidase papain, as well as the aspartic peptidase cathepsin D and the serine peptidase trypsin. Only cathepsin L was inhibited, while all of the other enzymes were not affected, even at high inhibitor concentrations. Cathepsin K rapidly cleaved the Tg-1



Scheme 1 Reaction mechanism proposed for cleavage of the Tg-1 domain by cathepsin L.

domain (see next section), while cathepsins B, S, X, and D, as well as papain and trypsin, did not. For cathepsin L, Eq. (1) was used to calculate υ_s, υ_z and λ (see Materials and methods). υ_s was a hyperbolic function of [Tg-1] and dropped to zero at saturating inhibitor concentrations, υ_z was independent of [Tg-1], and λ depended linearly upon [Tg-1] over the whole concentration range explored. The inhibitor affected K_m, but not the limiting rate V. This behaviour is characteristic of a fully competitive, slow-binding inhibition mechanism, in which the enzyme-inhibitor complex is formed in a single step (Morrison, 1982). Cleavage of the Tg-1 domain by cathepsin L was observed under second-order reaction conditions at high enzyme and Tg-1 concentrations (next section). The proposed complete reaction mechanism, including the substrate used to monitor reaction, is shown in Scheme 1.

I* collectively represents all possible fragments generated by proteolysis. Following the first attack on the most susceptible bond, two new peptides are generated and then smaller peptides are formed from these. If these are substrates for the enzyme, the system will consist of a mixture of substrates with different affinities. Thus, the path leading to I* contains multiple steps that remain kinetically inaccessible to experimental verification and is represented by the global constant k^*. Best-fit parameters and their standard errors, calculated by non-linear regression, were: $k_3=(8.94\pm0.35)\times10^6$ M^{-1} s^{-1}; $k_{-3}=(1.28\pm0.23)\times10^{-3}$ s^{-1}; and $K_i=k_{-3}/k_3=(1.4\pm0.3)\times10^{-10}$ M. There was no evidence that the Tg-1 domain was cleaved by proteolysis, thus behaving as a substrate, during assays with catalytic enzyme concentrations. In this case the steady-state slope had turned to a concave-upward profile for long reaction times. For calculation we used the concentration of the Tg-1 domain as determined spectrophotometrically. A titrated concentration of the Tg-1 domain was not available, since the only enzyme with which this would be possible is cathepsin L. Titration inevitably involves the use of both enzyme and inhibitor at relatively high concentrations, and under these conditions the inhibitor is degraded (see below). The high quality of the CD spectrum (Figure 1), which shows a remarkable difference between the native and the unfolded polypeptide, suggests that our testican Tg-1 domain preparation was properly folded and active. The inhibitory behaviour of the isolated Tg-1 domain towards cathepsin L is thus very similar to that of intact testican, with K_i values of 0.14 nM (this study) and 0.7 nM (Bocock et al., 2003). These results are in good agreement with our previous studies on equistatin (Lenarčič et al., 1997; Galeša et al., 2003) and the p41 fragment invariant chain (Bevec et al., 1996; Fiebiger et al., 2002), in which we demonstrated that the Tg-1 domains fully retain their functionality after excision from the rest of the parent molecule. Therefore, the testican Tg-1 domain may serve as a good model for further characterisation of the action of this type of inhibitor.

Degradation of Tg-1 domain of testican by cathepsins L and K

The kinetic measurements for inhibition of cathepsin L by testican Tg-1 domain were carried out with Tg-1 domain/cathepsin L molar ratios in the range 130:1–550:1 and picomolar enzyme concentrations. When molar ratios of 1:1–6:1 and concentrations of enzyme and inhibitor in the micromolar range were used, extensive degradation of the inhibitor molecule was observed. Reaction progress was followed by HPLC separation of the fragments, and degradation products were analysed by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF). A large number of fragments observed (Tables 1 and 2) were due to the presence of three differently expressed forms of testican Tg-1 domains and to partial reduction of the disulfide bridges during the process of cleavage. For the purposes of analysis we have omitted fragments that evidently arise from the minor E^{-4} and Gly^{1} forms. The time-dependent degradation of the domain enabled us to reveal the sequence of cleavage events. Within the tight cathepsin L/Tg-1 testican domain complex, the most susceptible bond appeared to be Gly^{22}–Ala^{23}. These amino acids are located at the bottom of the first loop in the close vicinity of the cathepsin L active site. After this primary cleavage, the first loop was attacked at two further sites: (i) between Gln^{12}-Lys^{13} and/or Asn^{9}-Arg^{10}, resulting in the excision of corresponding short peptides that occupy the active site (Lys^{13}-Gly^{22} and Arg^{10}-Gly^{22}, respectively); (ii) the bonds between Tyr^{34}-Lys^{35} and Lys^{35}-Ala^{36} became exposed at the same time, and their cleavage provoked the complete separation of the first loop from the rest of the molecule. Prolongation of the incubation time brought about only the release of disulfide-bonded fragments (Table 1).

The interaction of cathepsin K with the testican domain resulted in its rapid cleavage (Table 2). No inhibition of the enzyme could be detected, although it cannot be excluded that, during the interaction, the domain may enter the active site of cathepsin K as thyropin-like inhibitors normally do, i.e., with all three loops (Gunčar et al., 1999). The most exposed region of the Tg-1domain to cathepsin K was again the extended first loop, where the main fragmentation took place (Ser^{15}-Lys^{16}, Leu^{21}-Gly^{22} and Gly^{22}-Ala^{23}). Prolongation of the incubation time resulted in additional cleavages between Gln^{12}-Lys^{13} and on the C-terminal part of the Tg-1 molecule (Ala^{57}-Gly^{58} and Ala^{64}-Val^{65}).

Cysteine peptidases of the papain family have been shown to have primary substrate specificity at the P_2

Table 1 Peptides recovered after different times of incubation of testican Tg-1 with cathepsin L.

1 min		10 min		30 min		70 min	
Mass (Da)	Peptide	Mass (Da)	Peptide	Mass (Da)	Peptide	Mass (Da)	Peptide
5301.7±0.02	A^{23}-E^{70}	5306.7±0.11	A^{23}-E^{70}	5302.9±0.04	A^{23}-E^{70}	5302.6±0.03	A^{23}-E^{70}
2614.4±0.04	E-^{2}-G^{22}	4189.9±0.07	E-^{2}-K^{35}	3727.7±0.03	A^{36}-E^{70}	3727.3±0.02	A^{36}-E^{70}
3727.4±0.02	A^{36}-E^{70}	4060.3±0.09	E-^{2}-Y^{34}	2792.4±0.04	E-^{2}-N^{9}+A^{23}-K^{35}	2792.1±0.03	E-^{2}-N^{9}+A^{23}-K^{35}
1602.9±0.01	E-^{2}-Q^{12}	3855.6±0.03	K^{35}-E^{70}	2664.2±0.04	E-^{2}-N^{9}+A^{23}-Y^{34}	2664.0±0.03	E-^{2}-N^{9}+A^{23}-Y^{34}
		3727.4±0.02	A^{36}-E^{70}	2573.7±0.02	E-^{2}-N^{9}+I^{25}-K^{35}	2571.9±0.05	E-^{2}-N^{9}+I^{25}-K^{35}
		3189.4±0.03	E-^{2}-Q^{12}+A^{23}-K^{35}	2446.2±0.05	E-^{2}-N^{9}+I^{25}-Y^{34}	2445.9±0.04	E-^{2}-N^{9}+I^{25}-Y^{34}
		2792.2±0.04	E-^{2}-N^{9}+A^{23}-K^{35}	1461.6±0.07	A^{23}-Y^{34}	1461.6±0.06	A^{23}-Y^{34}
		2664.1±0.04	E-^{2}-N^{9}+A^{23}-Y^{34}	1589.9±0.08	A^{23}-K^{35}	1371.6±0.07	I^{25}-K^{35}
		1589.7±0.06	A^{23}-K^{35}	1371.7±0.07	I^{25}-K^{35}	1243.4±0.07	I^{25}-Y^{34}
		1461.6±0.07	A^{23}-Y^{34}	1243.6±0.09	I^{25}-Y^{34}		

The peptide molecular masses were determined by MALDI-TOF spectrometry and peptide fragments matching the masses were calculated with the PAWS program. For clarity, we show only the peptides derived from the major expressed form that starts with E-^{2} (region from -4 to -1 denotes amino acids that were not removed during processing in *P. pastoris*). Peptides connected by disulfide bonds are marked with "+". The cleavage points on the testican Tg-1 domain are presented on the sequence below. Bold black arrows represent the first occurring, most susceptible cleavages, black arrows the intermediate, and the grey arrows the weakest susceptible bonds. Letters shaded with grey (numbering starts from the left) denote the positions of loops. The positions of disulfide bridges are shown below the sequence.

position and to have a preference for hydrophobic amino acids (McGrath, 1999; Turk et al., 1998). Despite structural similarities, cathepsins L and K display differences in their binding preferences for P_2 residues. Cathepsin L favours both aromatic and aliphatic amino acid residues (Phe, Tyr, Trp, Val, Leu, Ile), while cathepsin K favours aliphatic amino acids (Ile, Leu, Val) and is unique in positioning Pro in the S_2 subsite, a property responsible for its potent intrahelical collagenolytic activity (Garnero et al., 1998; Lecaille et al., 2002). Our present results show that the amino acid residues at P_2 subsites of the identified cleavage points of the testican Tg-1 domain reflect the specificity of the cathepsin considered (Tables 1 and 2). Cathepsin L shows broader selection of amino acids at the P_2 position by utilising Tyr, Leu, Ile and Val, and can also employ Met and Ala, while cathepsin K is limited to Leu and Ile, with the one exception for Gly.

Evidence from experiments conducted under different conditions suggests that the dual nature of the Tg-1 testican domain as substrate and inhibitor of cathepsin L is due to the very low rate of degradation of the EI complex according to Scheme 1. An estimate for the rate of degradation of the Tg-1 domain was obtained from the time-dependent degradation of the inhibitor after incubating 0.39 μM cathepsin L with 2.3 μM Tg-1 domain at 24°C and pH 5.50 (data not shown). The final degradation

Table 2 Peptides recovered after different times of incubation of testican Tg-1 with cathepsin K.

10 min		30 min		60 min	
Mass (Da)	Peptide	Mass (Da)	Peptide	Mass (Da)	Peptide
5364.4±0.08	G^{22}-E^{70}	7239.5±0.01	E^{-2}-S^{15}+K^{16}-A^{64}	7293.2±0.06	E-^{2}-S^{15}+G^{22}-E^{70}
2614.1±0.03	E-^{2}-G^{22}	5364.2±0.08	G^{22}-E^{70}	7240.6±0.02	E-^{2}-S^{15}+K^{16}-A^{64}
2557.1±0.03	E-^{2}-L^{21}	5307.1±0.05	K^{16}-A^{64}	6964.7±0.06	E-^{2}-Q^{12}+G^{22}-E^{70}
1930.9±0.05	E-^{2}-S^{15}	2614.1±0.03	E-^{2}-G^{22}	6912.7±0.03	E-^{2}-Q^{12}+K^{16}-A^{64}
		1930.9±0.05	E-^{2}-S^{15}	5306.9±0.06	K^{16}-A^{64}
		1602.7±0.06	E-^{2}-Q^{12}	1930.8±0.05	E-^{2}-S^{15}
		1379.6±0.07	G^{58}-E^{70}	1602.6±0.06	E-^{2}-Q^{12}
				1379.5±0.07	G^{58}-E^{70}

The peptide molecular masses were determined by MALDI-TOF spectrometry and peptide fragments matching the masses were calculated with the PAWS program. For clarity, we show only the peptides derived from the major expressed form that starts with E-^{2} (region from -4 to -1 denotes amino acids that were not removed during processing in *P. pastoris*). Peptides connected by disulfide bonds are marked with "+". The cleavage points on the testican Tg-1 domain are presented on the sequence below. Symbols are explained in the legend of table 1.

products of the Tg-1 domain were no longer inhibitory, as ascertained by incubating them with cathepsin L and measuring the residual activity with Z-Phe-Arg-MCA (MCA, 4-methyl-7-coumarylamide) as substrate. Since evaluation of all the individual constants contributing to k^* in Scheme 1 was impossible, quantification was carried out by measuring the decreasing area of the HPLC peak corresponding to the Tg-1 domain and reaction profiles were analysed by the integrated Michaelis-Menten equation. Calculations afforded a first-order rate constant of 1.33×10^{-4} s^{-1} (half-life=5200 s), equivalent to V/K_m for the most susceptible bond of the Tg-1 domain, i.e., the first one cleaved by cathepsin L. Considering this result and the enzyme concentration used (0.39 μM), a first-order rate constant of 7.5×10^{-9} s^{-1} (half-life=9.2×10^{7} s) can be calculated for an enzyme concentration of 22 pM, i.e., the concentration used for measuring the inhibition constant (see the preceding section). Hence, the rate of Tg-1 domain degradation at catalytic enzyme concentration is approximately 18 000-fold lower than at high enzyme concentrations. Degradation is virtually invisible, and the Tg-1 domain behaves as an inhibitor. Cathepsin L used for the experiments shown in this paper was prepared in *E. coli* from inclusion bodies and was obtained by renaturation, thus making the presence of contaminating bacterial peptidases very unlikely. In order to rule out the possibility that Tg-1 domain degradation was not due to cathepsin L but to a contaminant enzyme, we repeated the cleavage experiment and fragment analysis using the same Tg-1 domain preparation and cathepsin L from two additional sources, namely the natural enzyme purified from human kidney and a recombinant form from *Pichia pastoris*. With all of the three different enzyme preparations we obtained the same cleavage products (data not shown).

Structural analysis of models

To explain the experimentally observed specific interaction of the Tg-1 domain of testican with cathepsin L, several models of their complex were built and compared to the models of complexes of the domain with cathepsin K and papain. The models were derived by a homology modelling procedure with amino acid sequence identity between the Tg-1 domain of testican and the template structure (p41 fragment) being above 30%. The major differences between the two molecules are present in the first subdomain and are a consequence of the insertion of five amino acids in the testican Tg-1 domain compared to the p41 fragment (Figure 2A). In the modelled domain, the extended α-helical region contributes to hydrophobic packing, and the region following the α-helix forms an enlarged loop that may have a more flexible conformation, similar to the corresponding loop in the p41 fragment, as observed from its NMR structure (Chiva et al., 2003). The interactions between the testican Tg-1 domain and cathepsin L in the modelled complex are electrostatic, polar and hydrophobic. Lys^{18} interacts electrostatically with Asp^{71}E and Asp^{114}E, and Arg^{27} and Lys^{35} with Asp^{137}E, Asp^{160}E and Asp^{162}E, the latter three forming the S_1' binding site (the letter E denotes enzyme residues). Polar interactions occur between Ser^{42} and both Cys^{25}E and Gln^{21}E, the latter representing the oxyanion hole. Hydrophobic interactions predominantly involve Leu^{21}, which fits into the deep and hydrophobic S_2 binding pocket of cathepsin L, and are mainly confined to the first and second loops of the domain (Figure 2B). These interactions are similar to interactions observed in the crystal structure of the p41 fragment complex with cathepsin L. Models of complexes of the domain with cathepsin K and papain (Figure 2C,D, respectively) show the same general binding mode, although the majority of favourable electrostatic interactions observed in the cathepsin L complex are not present. These observations are consistent with the finding that neither cathepsin K nor papain was inhibited by the Tg-1 domain. Furthermore, the surface of papain near the domain in the modelled complexes is positively charged, which may prevent binding of the domain to the enzyme (Figure 2D). These results lead to the conclusion that the domain binds to both cathepsin L and K, but not to papain.

It is still not possible to establish whether proteolytic cleavage of the domain by cathepsins L and K is a consequence of the geometry of the prolonged first loop or of the amino acid composition at the P_4-P_1/P_1' site, or maybe a combination of both. For example, testican and the first domain of equistatin exhibit a very similar arrangement of amino acid residues at P_4-P_1/P_1' positions within the first loop, i.e., Ser-Leu-Leu-Gly/Ala and Gly-Leu-Ile-Gly/Ala, respectively (Lenarčič et al., 1997). Only testican is cleaved by cathepsin L, which points to the position of the extended loop as a decisive factor. This could be explained by the more substrate-like nature of the extended first loop of the testican Tg-1 domain compared to the corresponding loop of the first domain of equistatin, which presumably is tighter and similar to the corresponding loop in the p41 fragment. However, another prerequisite for proteolytic cleavage is the accessibility of the scissile bond for nucleophilic attack of the sulfur atom of the active Cys residue (Figure 2E,F). This accessibility is in turn governed by the residues surrounding the putative cleavage site and their conformation. Finally, it should be noted that the interacting region of the testican Tg-1 domain is clearly different from a classical substrate, since it is not linear but rather composed of two parts, one of them interacting with S binding subsites and the other with S′ binding subsites. Nevertheless, the modelling results suggest that binding of the domain to the enzyme in an orientation similar to that found in the cathepsin L-p41 fragment complex is a prerequisite for proteolytic cleavage of the domain, since this is the only orientation that permits close contact between the domain and the enzyme.

Physiological aspects

There are several reports about the proteolytic processing of molecules that contain Tg-1 domains. Two C-terminal testican fragments (cleavage within the Tg-1 domain at position Tyr^{343}-Lys^{344} and Arg^{369}-Lys^{370}) with attached chondroitin sulfate chains were found as the major proteoglycans in human seminal plasma (Bonnet et al., 1992). The Tyr^{343}-Lys^{344} bond (testican unprocessed precursor counting) corresponds to the Tyr^{34}-Lys^{35} bond (Table 1), which is one of the cleavage points

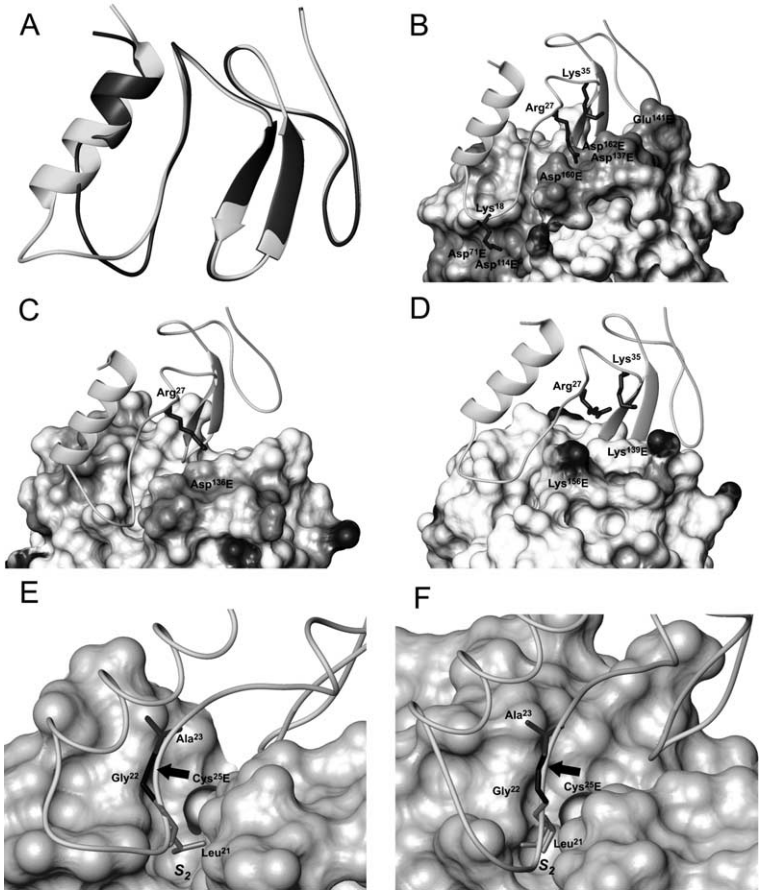


Figure 2 Three-dimensional models of testican Tg-domain in complex with target enzymes.
(A) Superimposition of the testican Tg-1 domain model (*grey*) and the p41 fragment (*black*). The insertion of five amino acids in the testican domain, as compared to the p41 fragment, results in an extended α-helix and altered geometry of the first loop. (B–D) Models of complexes of testican Tg d-1 domain and cathepsin L (B), cathepsin K (C) and papain (D). The surface of enzymes is shaded to represent the electrostatic potential: negatively charged residues are shown in grey and positively charged residues in black. Positively charged amino acid residues of the testican domain involved in favourable (cathepsins L and K) or unfavourable (papain) electrostatic interactions are shown as black sticks. (E,F) Testican Tg-domain displayed over the surfaces of the active sites of cathepsins L (E) and K (F). The position of the scissile bond (marked with an arrow) in the testican Tg d-1 domain relative to the active site cysteine residue ($Cys^{25}E$) is similar for both cathepsin L and K. In this substrate-like orientation, the P_2 residue (Leu^{21}) fits into the hydrophobic S_2 binding pocket of the enzyme. Images were generated using the program MOLMOL (Koradi et al., 1996).

detected after the interaction of the domain with cathepsin L. Interestingly, the Tg-1 domain was also proposed as a target cleavage region during the conversion of large plasma testican into smaller stable forms, although it was shown that processing can be blocked by certain serine peptidase inhibitors (BaSalamah et al., 2001). Nevertheless, in the present work we demonstrated that the Tg-1 domain of testican is very susceptible to cleavage by the two cysteine peptidases cathepsin L and cathepsin K. Therefore, the role of the testican Tg-1 domain must be different from that of the Tg-1 domain of the major histocompatibility complex class II associated invariant chain (p41 Ii), the processing of which has been studied in more detail (Fiebiger et al., 2002). Two key enzymes are crucial for processing p41 Ii, cathepsins S and L (Lennon-Dumenil et al., 2002) and, from the point of view of the Tg-1 domain, it remains intact after proteolysis and can serve either as an inhibitor of cathepsin L activity or as a chaperone in maintaining the pool of mature enzyme outside the cell (Fiebiger et al., 2002). The major difference between the two domains is the insertion of five amino acids on the N-terminal part of the testican domain, which probably has a great impact on the geometry and position of the first loop and consequently on the susceptibility to proteolytic attack.

In the light of the results reported here, testican may behave as an inhibitor or as a substrate of cathepsin L, depending on the relative concentrations available. In a close pericellular environment, where cathepsin L is likely to reach considerable local concentrations, proteolytic fragmentation of testican can be expected. Conversely, at a certain distance from a cathepsin L-releasing cell, the enzyme concentration is likely to be sufficiently low with respect to the proteoglycan to make proteolysis insignificant relative to inhibition. Further studies will be needed to focus on the possible physiological role of Tg-1 domains in the peri/extracellular space, where several cysteine peptidases are known to play an important role in remodelling of extracellular matrix proteins (McGrath, 1999). A number of modularly organised proteins with diverse functions, such as testican, nidogen, thyroglobulin, Trop and SMOC that contain one or more copies of Tg-1 domains, are considered to be extracellular proteins.

Materials and methods

Materials

Materials for cloning, expression and purification were obtained from the following sources: polymerase chain reaction (PCR) primers from MWG Biotech (Ebersberg, Germany); restriction enzymes, DNA ligase and deoxynucleotide from Promega (Madison, USA); Taq polymerase from Fermentas (Vilnius, Lithuania); salts and media for *P. pastoris* from Sigma-Aldrich (St. Louis, MO, USA); *P. pastoris* expression kit, pPIC9 plasmid and *P. pastoris* strain GS115 from Invitrogen (Carlsbad, CA, USA); Phenyl Sepharose, Sephadex G-50 and SP Sepharose FF from Amersham Pharmacia Biotech (Uppsala, Sweden); and Z-Phe-Arg-4-AMC, Z-Phe-Val-Arg-AMC, Z-Arg-Arg-AMC and Z-Phe-Arg-pNA from Bachem AG (Bubendorf, Switzerland). 2,4-Dinitrophenyl-Gly-Phe-Phe-Trp-OH was synthesised according to Klemenčič et al (2000) and the cDNA coding for the testican protein was kindly provided by C.J.S. Edgell (University of North Carolina, USA). Papain was prepared as described in Blumberg et al. (1970), human cathepsins B, D, L, K, S as in Kuhelj et al. (1995), Kregar et al. (1977), Dolinar et al. (1995), Hwang and Chung (2002) and Brömme et al., (1993), respectively, and cathepsin X according to Klemenčič et al. (2000). Bovine trypsin was obtained from Sigma-Aldrich.

Cloning and expression of testican Tg-1 domain

The DNA segment coding for the Tg-1 domain of testican (residues Gly^{310}–Glu^{379}) was amplified with the primers 5′-GGA TCC CTC GAG AAA AGA GAG GCT GAA GCT GGT CTC CCT TGC CAG AAT GAA ATG AAC AGA ATT-3′ (sense), and 5′-AAT TCG CGG CCG CTT ACT CCT CTT CAC AGC T-3′ (antisense) using a plasmid encoding full-length testican as a template. The amplified DNA was digested with *Xho*I and *Not*I restriction endonucleases and ligated into pPIC9 digested with the same enzymes. The recombinant construct was amplified in *Escherichia coli* strain DH5α and the correct construction of the plasmid was confirmed by restriction analysis and DNA sequencing using an ABI PRISM 310 DNA sequencer (Perkin Elmer, Wellesley, MA, USA). After linearisation of the pPIC9 construct with *Sal*I, the *Pichia pastoris* strain GS115 was transformed by electroporation with Gene Pulser (Bio-Rad, Hercules, CA, USA) according to the manufacturer's protocol. The transformed cells were plated on MD agar plates [0.34% yeast nitrogen base (YNB), 1% ammonium sulfate, 400 μg/l biotin, 2% D-glucose, 1.5% agar]. Resulting colonies were used to inoculate 10 ml of BMGY medium (1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 6.0, 0.34% YNB, 1% ammonium sulfate, 400 μg/l biotin, 1% glycerol) and grown overnight at 30°C in a shaking incubator at 250 rpm. Cells were collected by centrifugation (2500 *g* for 5 min at 23°C) and gently resuspended in 10 ml of BMMY medium (same as BMGY medium but containing 0.5% methanol instead of 1% glycerol). Cultures were incubated for an additional 96 h with daily supplements of 50 μl of methanol. Clone production was monitored by electrophoretic analysis of the supernatants on 12% polyacrylamide SDS gels (Laemmli, 1970). The colony with the highest expression level of testican Tg-1 domain was used to prepare protein in larger quantities, using 2-l shake-flasks under the conditions specified above. The cells were pelleted and the supernatant was stored at -20°C.

Purification of the Tg-1 domain

Ammonium sulfate was added to culture supernatants to a final concentration of 1.0 M and the pH was adjusted to 6.0. The protein solution was loaded onto a phenyl-Sepharose column equilibrated with 0.1 M phosphate buffer containing 1.0 M ammonium sulfate. Proteins were eluted with a linear gradient of ammonium sulfate (from 1 to 0 M) in 0.1 M phosphate buffer, pH 6.0 and analysed by SDS-PAGE, in which the Tg-1 domain of testican migrated as a band with a molecular mass of approximately 8 kDa. Fractions containing testican domain were further purified on a Sephadex G-50 column equilibrated with 0.01 M Tris/HCl buffer, pH 7.3 in 0.1 M NaCl. Final purification of the testican domain was achieved by ion-exchange chromatography on SP Sepharose FF pre-equilibrated with 0.01 M acetate buffer, pH 4.5. Proteins were eluted with a linear gradient of NaCl (0–0.5 M) in the same buffer. The concentration of the Tg-1 domain was determined spectrophotometrically using an absorption coefficient of 9890 M^{-1} cm^{-1} at 280 nm.

N-Terminal sequence determination

N-Terminal sequencing was carried out by automated Edman degradation using an Applied Biosystems liquid pulse sequencer 475A (Foster City, CA, USA) connected on line to a phenylthiohydantoin analyser 120A from the same manufacturer.

Circular dichroism spectroscopy

The far-UV CD spectrum of recombinant Tg-1 domain of testican was measured at 4°C on an Aviv 60DS spectrometer (Piscataway, NJ, USA) using a quartz cell with 1-mm path length. In the wavelength region 250–185 nm, data points were recorded at 1-nm intervals with a dwell time of 10 s. Protein was dissolved at 0.1 mg/ml in 10 mM sodium phosphate buffer, pH 5.6. For thermal denaturation, the sample was exposed to 88°C for 1 h in the presence of 12.5 μM dithiothreitol (DTT) and the spectra were recorded at 88°C. To confirm the absence of turbidity, the absorbance of protein solutions was scanned from 350 to 240 nm on a Lambda 18 spectrophotometer (Perkin Elmer) prior to CD measurement.

Kinetic measurements

Enzymes were assayed following standard procedures using the substrate Z-Phe-Arg-MCA for papain, cathepsins L and cathepsin K, Z-Arg-Arg-MCA for cathepsin B, Z-Phe-Val-Arg-MCA for cathepsin S, and 2,4-dinitrophenyl-Gly-Phe-Phe-Trp-OH for cathepsin X. Inhibitory effects of the testican Tg-1 domain were studied by incubating it with an enzyme for 30 min at 25°C before adding the substrate and measuring the residual activity.

The kinetic inhibition mechanism of cathepsin L by the Tg-1 domain was investigated under pseudo-first-order conditions essentially as described for the equistatin Tg-1 type inhibitor (Lenarčič et al., 1997). Testican Tg-1 domain (2.9–12 nM final concentration) and the fluorogenic substrate Z-Phe-Arg-MCA (10 μM final concentration) were dissolved in 1.97 ml of 0.1 M sodium acetate buffer, pH 5.50 containing 1 mM EDTA, at 24°C. The reaction was started by the addition of 30 μl of cathepsin L, which was pre-activated with DTT, giving a final enzyme concentration of 22 pM. Progress curves were recorded for 30–60 min on a C-61 fluorimeter (Photon Technology International, Lawrenceville, GA, USA) and the conditions chosen assured less than 8–10% substrate consumption during each experiment. Primary data consisted of an exponential rise in product concentration followed by a linear steady state, a typical characteristic of slow-binding inhibition (Morrison, 1982). The general, integrated rate equation [Eq. (1)] was fitted to data to obtain best fits of the parameters v_s (velocity at steady state), v_z (velocity at time zero), and λ (apparent pseudo-first-order rate constant for exponential approach to a steady state).

Figure 3 Amino acid sequence alignment of the p41 fragment and testican Tg-1 domain.
Alignment was carried out with the algorithm of the Modeler program (Šali and Blundell, 1993). Identical amino acids are indicated by shaded boxes. The position of inserted amino acids is underlined. Numbers above the sequences denote the positions of loops.

$$[P]=\upsilon_s t+(\upsilon_z-\upsilon_s)(1-e^{-\lambda t})/\lambda \quad (1)$$

υ_s, υ_z and λ were then plotted against the Tg-1 concentration to determine the inhibition mechanism.

The Selwyn (1965) method, which consists of plotting progress curves as product concentration vs. enzyme concentration multiplied by time, was used to check any loss of enzyme activity by denaturation, adherence to the cuvette walls or other causes during the relatively long incubation required by the experiments. There was no enzyme activity loss for measurement times of up to at least 20 min, so that data collected in this time range were selected for calculations.

Proteolytic cleavage of Tg-1 domain of testican

The proteolytic degradation of testican Tg-1 domain (2.1–3.5 μM final concentration) by activated cathepsin L (0.39 μM final concentration) or cathepsin K (40 nM final concentration) was carried out in 0.1 M sodium acetate buffer, pH 5.50, containing 1 mM EDTA. The mixtures were incubated for different periods of time at 25°C. The progress of degradation was followed on an reverse-phase HPLC system (HP 1100, Hewlett-Packard, Palo Alto, CA, USA) using a CromSpher C18 (Varian, Palo Alto, CA, USA) column equilibrated with 0.1% (v/v) trifluoroacetic acid in water. Elution of peptides was carried out with 90% (v/v) acetonitrile containing 0.1% (v/v) trifluoroacetic acid. In control experiments, testican Tg-1 and cathepsins L and K were separately incubated under the same conditions for 60 min at 37°C. Progress curves for Tg-1 domain degradation by cathepsin L were analysed using the integrated Michaelis-Menten equation as described by Duggleby (2001).

Identification of cleavage sites by MALDI-TOF

Samples for identification of cleaved bonds were prepared as described under Proteolytic cleavage of Tg-1 domain of testican. Mixtures of peptides were desalted on a reverse-phase HPLC system with a short run (2 min) in an acetonitrile gradient and analysed by MALDI-TOF. Desalted peptides were dried and dissolved in a saturated solution of α-cyano-4-hydroxycinnamic acid in 0.1% trifluoroacetic acid (v/v), 60% (v/v) acetonitrile and placed on the MALDI target. MALDI data were recorded on a Bruker Biflex III (Billerica, MA, USA) instrument equipped with a scout ion source. Spectra were acquired with pulsed ion extraction in reflectron mode using a 337-nm nitrogen laser.

Homology modelling

The models of complexes of testican Tg-1 domain with papain and cathepsins K and L were calculated with the homology modelling program Modeler (Šali and Blundell, 1993) using the crystal structures of papain (Kamphuis et al., 1984), cathepsin K (Zhao et al., 1997) and p41 fragment complexed with cathepsin L (Gunčar et al., 1999) as templates (PDB codes 9PAP, 1ATK and 1ICF, respectively). Prior to modelling, the template structures were cleaned of water molecules and non-native atoms (oxygen atoms on oxidised Cys^{25} in papain and E-64 in the cathepsin K structure). In the initial modelling stage, the alignments of the Tg domains (Figure 3) and of the peptidases were used, along with template structures to build 20 models of each complex. During this stage the segment of the Tg-1 domain including residues Cys^4–Leu^{14} was restrained to an α-helical conformation based on predictions by the PsiPred program (McGuffin et al., 2000). For each complex, one model was chosen from the pool of the initial 20 and the segment Ser^{15}–Ala^{23} of the domain was subjected to a long simulated annealing refinement protocol implemented in Modeler. Of the 20 refined models, five models with the lowest value of the Modeler objective function were checked with PROCHECK (Laskowski et al., 1996) and analysed.

Acknowledgments

The financial support of the Ministry of Education, Science, and Sport is acknowledged. We also thank Mojca Trstenjak Prebanda, Ivica Klemenčič and Marko Mihelič for providing cathepsins K, X and S, respectively, and Roger Pain for helpful discussion and critical reading of the manuscript.

References

Baici, A., Lang, A., Hörler, D., Kissling, R., and Merlin, C. (1995). Cathepsin B in osteoarthritis: zonal variation of enzyme activity in human femoral head cartilage. Ann. Rheum. Dis. *54*, 289–297.

BaSalamah, M.A., Marr, H.S., Duncan, A.W., and Edgell, C.J. (2001). Testican in human blood. Biochem. Biophys. Res. Commun. *283*, 1083–1090.

Bevec, T., Stoka, V., Pungerčič, G., Dolenc, I., and Turk, V. (1996). Major histocompatibility complex class II-associated p41 invariant chain fragment is a strong inhibitor of lysosomal cathepsin L. J. Exp. Med. *183*, 1331–1338.

Blumberg, S., Schechter, I., and Berger, A. (1970). The purification of papain by affinity chromatography. Eur. J. Biochem. *15*, 97–102.

Bocock, J.P., Edgell, C.J., Marr, H.S., and Erickson, A.H. (2003). Human proteoglycan testican-1 inhibits the lysosomal cysteine protease cathepsin L. Eur. J. Biochem. *270*, 4008–4015.

Bonnet, F., Perin, J.P., Maillet, P., Jolles, P., and Alliel, P.M. (1992). Characterization of a human seminal plasma glycosaminoglycan-bearing polypeptide. Biochem. J. *288*, 565–569.

Bonnet, F., Perin, J.P., Charbonnier, F., Camuzat, A., Roussel, G., Nussbaum, J.L., and Alliel, P.M. (1996). Structure and cellular distribution of mouse brain testican. Association with the postsynaptic area of hippocampus pyramidal cells. J. Biol. Chem. *271*, 4373–4380.

Brake, A.J., Merryweather, J.P., Coit, D.G., Heberlein, U.A., Masiarz, F.R., Mullenbach, G.T., Urdea, M.S., Valenzuela, P., and Barr, P.J. (1984). α-Factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae.* Proc. Natl. Acad. Sci. USA *81*, 4642–4646.

Brömme, D., Bonneau, P.R., Lachance, P., Wiederanders, B., Kirschke, H., Peters, C., Thomas, D.Y., Storer, A.C., and Vernet, T. (1993). Functional expression of human cathepsin S in *Saccharomyces cerevisiae*. Purification and characterization of the recombinant enzyme. J. Biol. Chem. *268*, 4832–4838.

Cereghino, J.L. and Cregg, J.M. (2000). Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol. Rev. *24*, 45–66.

Chapman, H.A., Riese, R.J., and Shi, G.P. (1997). Emerging roles for cysteine proteases in human biology. Annu. Rev. Physiol. *59*, 63–88.

Chiva, C., Barthe, P., Codina, A., Gairi, M., Molina, F., Granier, C., Pugniere, M., Inui, T., Nishio, H., Nishiuchi, Y., et al. (2003). Synthesis and NMR structure of p41icf, a potent inhibitor of human cathepsin L. J. Am. Chem. Soc. *125*, 1508–1517.

Dolinar, M., Maganja, D.B., and Turk, V. (1995). Expression of full-length human procathepsin L cDNA in *Escherichia coli* and refolding of the expression product. Biol. Chem. Hoppe-Seyler *376*, 385–388.

Duggleby, R.G. (2001). Quantitative analysis of the time courses of enzyme-catalyzed reactions. Methods *24*, 168–174.

Fiebiger, E., Maehr, R., Villadangos, J., Weber, E., Erickson, A., Bikoff, E., Ploegh, H.L., and Lennon-Dumenil, A.M. (2002). Invariant chain controls the activity of extracellular cathepsin L. J. Exp. Med. *196*, 1263–1269.

Galeša, K., Pain, R., Jongsma, M.A., Turk, V., and Lenarčič, B. (2003). Structural characterization of thyroglobulin type-1 domains of equistatin. FEBS Lett. *539*, 120–124.

Garnero, P., Borel, O., Byrjalsen, I., Ferreras, M., Drake, F.H., McQueney, M.S., Foged, N.T., Delmas, P.D., and Delaisse, J.M. (1998). The collagenolytic activity of cathepsin K is unique among mammalian proteinases. J. Biol. Chem. *273*, 32347–32352.

Gunčar, G., Pungerčič, G., Klemenčič, I., Turk, V., and Turk, D. (1999). Crystal structure of MHC class II-associated p41 Ii fragment bound to cathepsin L reveals the structural basis for differentiation between cathepsins L and S. EMBO J. *18*, 793–803.

Hartmann, U. and Maurer, P. (2001). Proteoglycans in the nervous system: the quest for functional roles *in vivo*. Matrix Biol. *20*, 23–35.

Hwang, H.S. and Chung, H.S. (2002). Preparation of active recombinant cathepsin K expressed in bacteria as inclusion body. Protein Express. Purif. *25*, 541–546.

Kamphuis, I.G., Kalk, K.H., Swarte, M.B., and Drenth, J. (1984). Structure of papain refined at 1.65 Å resolution. J. Mol. Biol. *179*, 233–256.

Kidrič, M., Fabian, H., Brzin, J., Popovič, T., and Pain, R.H. (2002). Folding, stability, and secondary structure of a new dimeric cysteine proteinase inhibitor. Biochem. Biophys. Res. Commun. *297*, 962–967.

Klemenčič, I., Carmona, A.K., Cezari, M.H., Juliano, M.A., Juliano, L., Gunčar, G., Turk, D., Križaj, I., Turk, V., and Turk, B. (2000). Biochemical characterization of human cathepsin X revealed that the enzyme is an exopeptidase, acting as carboxymonopeptidase or carboxydipeptidase. Eur. J. Biochem. *267*, 5404–5412.

Koepf, E.K., Petrassi, H.M., Sudol, M., and Kelly, J.W. (1999). An isolated three-stranded antiparallel β-sheet domain that unfolds and refolds reversibly; evidence for a structured hydrophobic cluster in urea and GdnHCl and a disordered thermal unfolded state. Protein Sci. *8*, 841–853.

Koradi, R., Billeter, M., and Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. J. Mol. Graph. *14*, 29–32, 51–55.

Kregar, I., Urh, I., Umezawa, H., and Turk, V. (1977). Purification of cathepsin D by affinity chromatography on pepstatin Sepharose column. Croat. Chem. Acta *49*, 587–592.

Kuhelj, R., Dolinar, M., Pungerčar, J., and Turk, V. (1995). The preparation of catalytically active human cathepsin B from its precursor expressed in *Escherichia coli* in the form of inclusion bodies. Eur. J. Biochem. *229*, 533–539.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680–685.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR *8*, 477–486.

Lecaille, F., Choe, Y., Brandt, W., Li, Z., Craik, C.S., and Brömme, D. (2002). Selective inhibition of the collagenolytic activity of human cathepsin K by altering its S2 subsite specificity. Biochemistry *41*, 8447–8454.

Lenarčič, B. and Bevec, T. (1998). Thyropins: new structurally related proteinase inhibitors. Biol. Chem. *379*, 105–111.

Lenarčič, B. and Turk, V. (1999). Thyroglobulin type-1 domains in equistatin inhibit both papain-like cysteine proteinases and cathepsin D. J. Biol. Chem. *274*, 563–566.

Lenarčič, B., Ritonja, A., Štrukelj, B., Turk, B., and Turk, V. (1997). Equistatin, a new inhibitor of cysteine proteinases from *Actinia equina*, is structurally related to thyroglobulin type-1 domain. J. Biol. Chem. *272*, 13899–13903.

Lenarčič, B., Krishnan, G., Borukhovich, R., Ruck, B., Turk, V., and Moczydlowski, E. (2000). Saxiphilin, a saxitoxin-binding protein with two thyroglobulin type 1 domains, is an inhibitor of papain-like cysteine proteinases. J. Biol. Chem. *275*, 15572–15577.

Lennon-Dumenil, A.M., Bakker, A.H., Wolf-Bryant, P., Ploegh, H.L., and Lagaudriere-Gesbert, C. (2002). A closer look at proteolysis and MHC-class-II-restricted antigen presentation. Curr. Opin. Immunol. *14*, 15–21.

McGrath, M.E. (1999). The lysosomal cysteine proteases. Annu. Rev. Biophys. Biomol. Struct. *28*, 181–204.

McGuffin, L.J., Bryson, K., and Jones, D.T. (2000). The PSIPRED protein structure prediction server. Bioinformatics *16*, 404–405.

Molina, F., Bouanani, M., Pau, B., and Granier, C. (1996). Characterization of the type-1 repeat from thyroglobulin, a cysteine-rich module found in proteins from different families. Eur. J. Biochem. *240*, 125–133.

Morrison, J.F. (1982). The slow-binding and slow, tight-binding inhibition of enzyme-catalysed reactions. Trends Biochem. Sci. *7*, 102–105.

Nagase, T., Seki, N., Ishikawa, K., Ohira, M., Kawarabayasi, Y., Ohara, O., Tanaka, A., Kotani, H., Miyajima, N., and Nomura, N. (1996). Prediction of the coding sequences of unidentified human genes. VI. The coding sequences of 80 new genes (KIAA0201–KIAA0280) deduced by analysis of cDNA clones from cell line KG-1 and brain. DNA Res. *3*, 321–329, 341–354.

Nakada, M., Yamada, A., Takino, T., Miyamori, H., Takahashi, T., Yamashita, J., and Sato, H. (2001). Testican 2 abrogates inhibition of membrane-type matrix metalloproteinases by other testican family proteins. Cancer Res. *61*, 8896–8902.

Selwyn, M.J. (1965). A simple test for inactivation of an enzyme during assay. Biochim. Biophys. Acta *105*, 193–195.

Šali, A. and Blundell, T.L. (1993). Comparative protein modeling by satisfaction of spatial restraints. J. Mol. Biol. *234*, 779–815.

Turk, D., Gunčar, G., Podobnik, M., and Turk, B. (1998). Revised definition of substrate binding sites of papain-like cysteine proteases. Biol. Chem. *379*, 137–147.

Vannahme, C., Schubel, S., Herud, M., Gosling, S., Hulsmann, H., Paulsson, M., Hartmann, U., and Maurer, P. (1999). Molecular cloning of testican-2: defining a novel calcium-binding proteoglycan family expressed in brain. J. Neurochem. *73*, 12–20.

Woody, R.W. (1994). Contributions of tryptophan side chains to the far-ultraviolet circular dichroism of proteins. Eur. Biophys. J. *23*, 253–262.

Yamashita, M. and Konagaya, S. (1996). A novel cysteine protease inhibitor of the egg of chum salmon, containing a cysteine-rich thyroglobulin-like motif. J. Biol. Chem. *271*, 1282–1284.

Zhao, B., Janson, C.A., Amegadzie, B.Y., D'Alessio, K., Griffin, C., Hanning, C.R., Jones, C., Kurdyla, J., McQueney, M., et al. (1997). Crystal structure of human osteoclast cathepsin K complex with E-64. Nat. Struct. Biol. *4*, 109–111.

Received July 7, 2004; accepted November 25, 2004